

Time Warping Algorithms and its Applications on Financial Time Series

摘要

DTWGA 在金融网络中的应用；DTW算法（含有更多的约束条件）在投资管理上的应用；利用信号检测理论和从模糊集中借鉴的概念合并人们使用的技术形态或者图表形态。

introduction

非线性的难以处理，例如 $S_1(t + \Delta t) = f(S_2(t))$ ，当 Δt 不等于一个常数时，难以找到相关性。所以不同的 **时间扭曲** 方法被提出解决这个问题。DTW 常用于语音识别，检测心率失常等；三个约束：边界条件；步长；单调性。时间复杂度为 $O(N^2)$ ，可以用来识别波峰、波谷。

SDT（贝叶斯信号检测理论）

GA 遗传算法

第一章解释DTWGA和DTW算法的细节

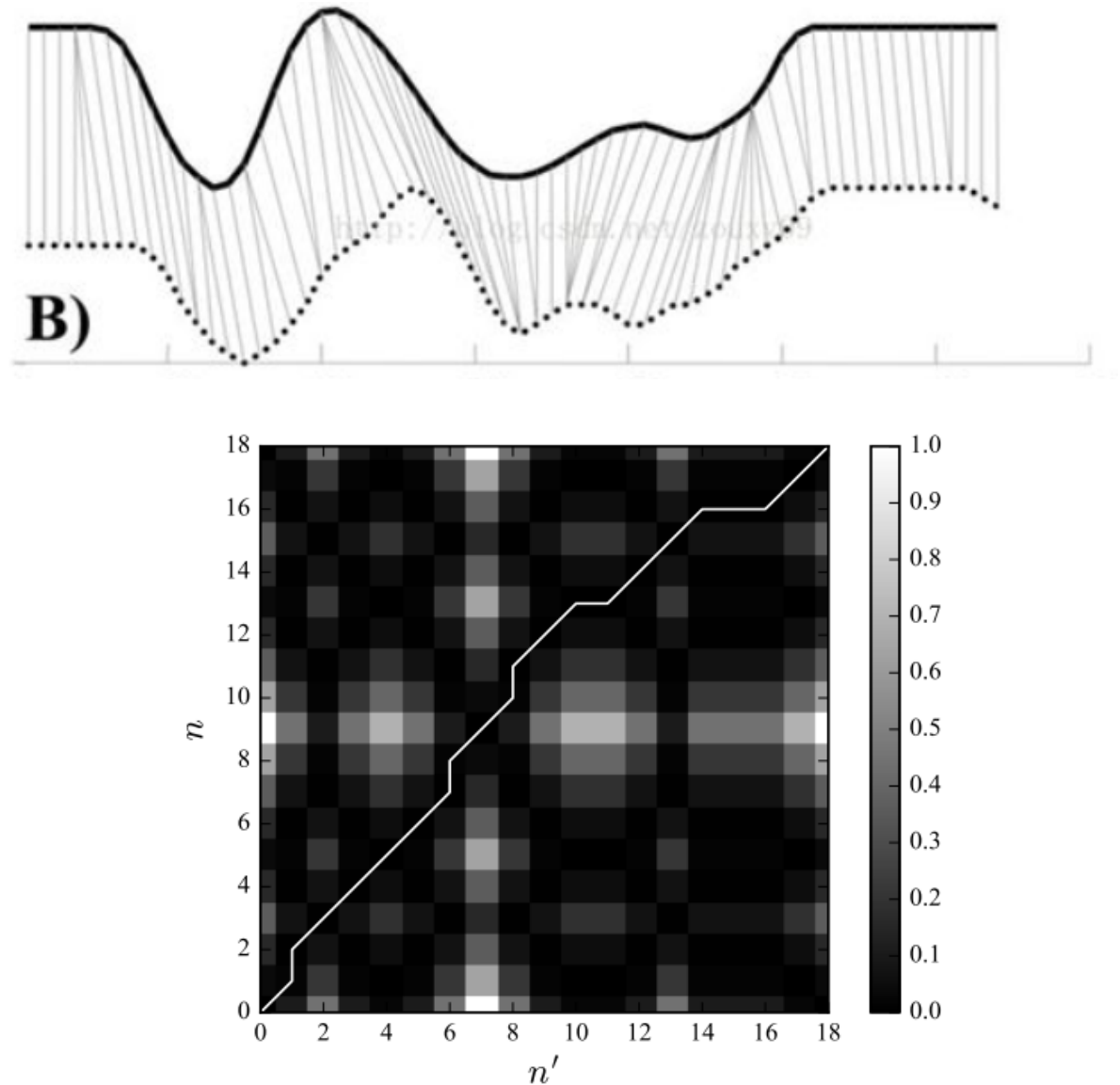
第二章贝叶斯信号检测理论、似然比和阈值估计，K-mean算法用于模式发现

第三章结合Parrondo玩具模型说明他在投资组合管理的影响；将模糊集概念加入到SDT中

D T W (动态时间规整)

DTW 可以计算两个时间序列的相似度，尤其适用于不同长度、不同节奏的时间序列（比如不同的人读同一个词的音频序列）。DTW 将自动warping扭曲 时间序列（即在时间轴上进行局部的缩放），使得两个序列的形态尽可能的一致，得到最大可能的相似度。

DTW采用了动态规划DP（dynamic programming）的方法来进行时间规整的计算，可以说，**动态规划方法在时间规整问题上的应用就是DTW。**



起始条件:

$$L_{min}(1,1) = M(1,1)$$

递推条件:

$$L_{min}(i,j) = \min\{L_{min}(i,j-1), L_{min}(i-1,j) + L_{min}(i-1,j-1)\} + M(i,j)$$

GA遗传算法

染色体：可行解（编码）

适应度矩阵：N个染色体对应的适应度（优劣程度）

选择概率矩阵：每个染色体的适应度除以总适应度之和，代表着被选中作为父母生成下一代的概率。

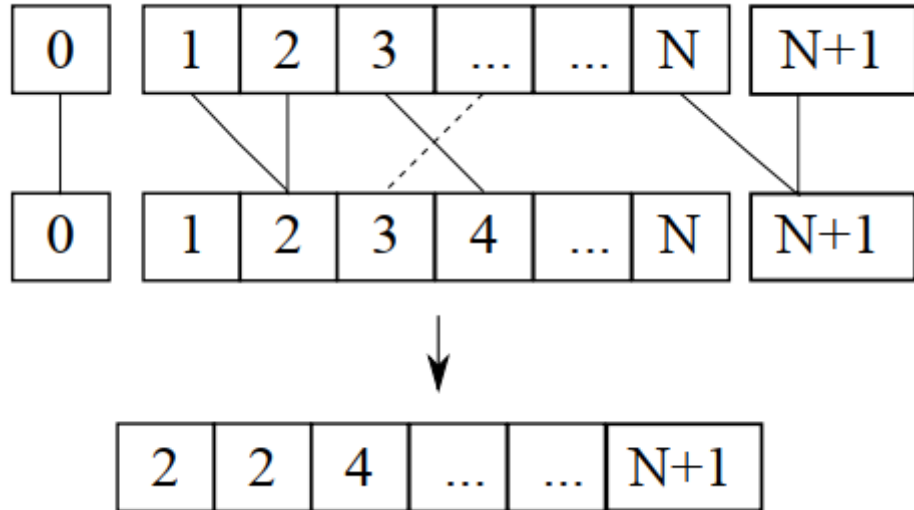
交叉：用父母的染色体生成子代。

变异：随机对基因进行变异

DTWGA

编码：

设有 N_p 条染色体(长度都为 N),第 k 条可以用 $C_k = \{\tau'_n | n = 1, 2, \dots, N\}$ 来表示。如下图即为 $\{2, 2, 4, \dots, N+1\}$



$$\begin{aligned} \tau_n &\leq \tau_{n+1}, \forall n \\ 0 &\leq \tau_n \leq N+1, \forall n \end{aligned}$$

突变：

适应度最好的一条染色体不动；其余 $(N_p - 1)$ 个染色体都随机变化 N_m 个基因，即共有 $(N_p - 1)N_m$ 个突变，若第 n 个基因发生突变，则其突变范围为 $[\tau_{n-1}, \tau_{n+1}]$ 。

适应度函数：

$$s_k = |f_0 - g_0| + |f_{N+1} - g_{N+1}| + \sum_{n=1}^N |f_n - g_{\tau_n}|$$

选择：

将上述 s_k 排序得： $s_1 \leq s_2 \leq \dots \leq s_k \leq s_{k+1} \leq s_{N_p}$ ，越小代表拟合度越好， N_p 代表该代的总染色体数目。

从中剔除 N_k 个最差的染色体，然后从剩下拟合度较好的 $N_p - N_k$ 中选择 N_k 个来替代。

其中下式越大的 k 越容易被选中：

$$P_p(k) = \frac{s_k}{\sum_{j=1}^{N_p - N_k} s_j}$$

迭代G次

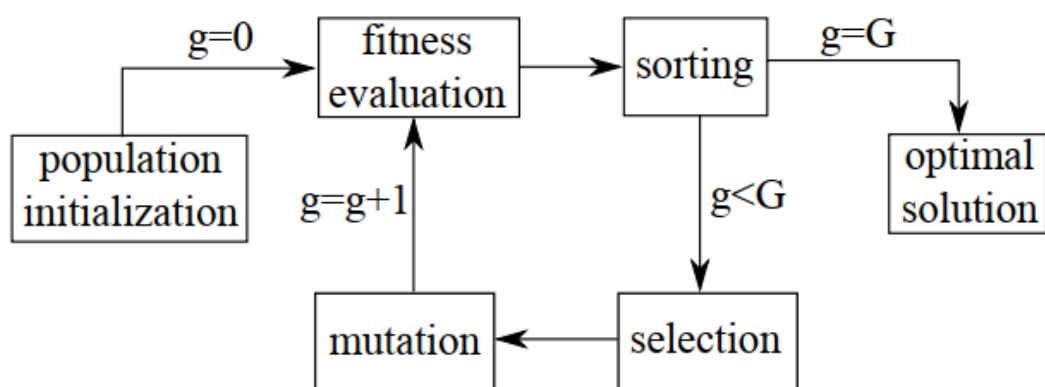


Figure 2.5: Flow chart of dTWGA.